

工作报告

一、调研

本周调研 map-matching (MM) 的具体研究情况，围绕 3 个问题调研：(1) 目前 MM 主要的方法有哪些 (2) 是否有人用 DL 或机器学习的方法解决 MM (3) 使用的数据集

- 1、目前主要的方法分为几何、拓扑、概率、高级技术
- 2、几何的方法又可以细分为利用点到点、点到线、线到线的距离来寻找对应的 road^[1]
- 3、拓扑的方法则是考虑轨迹和地图的拓扑信息，即利用轨迹、地图的点-线关系及道路的形状、邻接关系来找到 GPS 最合适的 road
- 4、基于概率的方法则认为 GPS 误差由两部分组成，其一是测量误差，它服从高斯分布，另一部分则是采样误差，即由于采样频率过低导致的轨迹不确定性^[3]
- 5、主流的高级方法大多基于两种理论：a、隐马尔科夫模型 (Hidden Markov Model, HMM) b、Kalman filter c、fuzzy logic (未调研)
- 6、隐马尔科夫模型，假设真实坐标为隐马尔科夫模型的状态，状态不可测；假设 GPS 坐标为隐马尔科夫模型的可测现象^[4]。再利用数据，去学习隐马尔科夫的状态转换概率矩阵和状态-现象概率矩阵。学习完这两个矩阵以后，已知一系列 GPS 坐标 (观测现象) 即可求得隐藏的真实状态 (实际位置)，从而解出 Map-Matching 问题。目前的研究则在于如何建立 MM 到 HMM 的关系，主要是如何定义实际位置的状态
- 7、Kalman filter 是用来解决测量误差的一种方法，首先构造一个模型，利用历史数据预测真实位置，再直接观测真实位置 (有测量误差)，最后预测位置和真实位置一起加权决定 map-matching 的结果
- 8、如果用 AI4Vis 来解决 MM 的优点是，a、前人没有用过该方法 b、它把不能用 dl 解决的问题转换到了 dl 能解决的问题 c、有监督的机器学习方法，能充分利用前人的经验。目前存在的困难是，具体落实还有很多细节没有考虑好。
- 9、MM 有关的数据集网上有很多，目前没用统一的数据集用于训练和验证。

二、与黄兆嵩博士的讨论

- 1、询问了 GPS 数据，地图数据，道路数据的具体组织方法。得到的结论是道路数据和轨迹数据是一样的，都是由一系列的点组成。
- 2、我对 map-matching 实现细节上还存在一些不清楚的地方，黄兆嵩建议我复现一个 MM 算法

三、与夏老师的讨论

- 1、我把我调研的论文情况和夏老师说了
- 2、目前我需要同时调研 MM 和 DL 方面的论文，不知道以哪个为主
- 3、目前 MM 方法中绑路困难的情况的示例如下，图中点 P1-7 代表 GPS 原始数据点。如图一所示当 P3 到 BC 和 BD 的距离相等时，则难以正确绑路。该情况有文献^[2]采用了数据融合的方法，即对比该点的速度数据和该路按时间段聚类的速度数据，对比该点的行驶方向来判断究竟 P3 与哪条路对应。图二正确的路径是 AC-CD，但由于 CD 和 FG 比较接近，测得的 GPS 坐标更接近于 FG，因此基于几何、拓扑的方法易把该轨迹对应到 AC-CF-FG 道路上

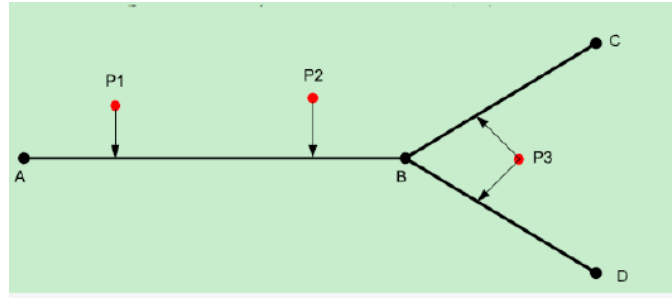
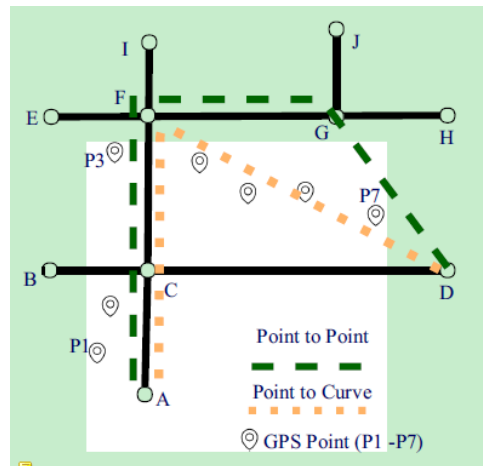


图 1



图二

4、夏老师认为是道路比较复杂，单个坐标不可信导致的。该问题与轨迹预测类似，而我则认为该问题虽与轨迹预测类似但实则有些不同，该问题比较类似于测量误差矫正的问题。我们没有达成共识，该问题需要进一步查阅资料去确认

5、夏老师认为，我对 MM 问题没有搞清楚，暂时先不管 DL 的研究，先去调研清楚 MM 的研究以及和道路预测的关系

6、重点调研目前 MM 上所使用的机器学习方法

参考文献

- [1] Singh, Jaiteg, et al. "Evaluating the performance of map matching algorithms for navigation systems: an empirical study." Spatial Information Research (2018): 1-12.
(注，该文章发表的杂志知名度不高，大致参考即可)
- [2] Hu, Gang, et al. "If-matching: Towards accurate map-matching with information fusion." IEEE Transactions on Knowledge and Data Engineering 29.1 (2017): 114-127.
- [3] Pfoser, Dieter, and Christian S. Jensen. "Capturing the uncertainty of moving-object representations." International Symposium on Spatial Databases. Springer, Berlin, Heidelberg, 1999. (500 多次引用)
- [4] Lou, Yin, et al. "Map-matching for low-sampling-rate GPS trajectories." Proceedings of the 17th ACM SIGSPATIAL international conference on advances in geographic information systems. ACM, 2009. (500 多次引用)

四、工作时间

周一-周四：9:00-21:00 共 $11 \times 5 = 55$ 小时

周五：8:50-12:00 共 3 小时

周五下午-周六 回合工大参加六级考试，给罗月童（我合工大的导师）老师组的学生做交流心得报告

周日：9:00-21:00 共 7 小时

本周共 65 小时